

Inference on the Cointegration Rank in Fractionally Integrated Processes

Jörg Breitung
Humboldt University Berlin

Uwe Hassler
Free University Berlin

This version: July 5, 2000

Abstract

For univariate time series we suggest a new variant of efficient score tests against fractional alternatives. This test has three important merits. First, by means of simulations we observe that it is superior in terms of size and power in some situations of practical interest. Second, it is easily understood and implemented as a slight modification of the Dickey-Fuller test, although our score test has a limiting normal distribution. Third and most important, our test generalizes to multivariate cointegration tests just as the Dickey-Fuller test does. Thus it allows to determine the cointegration rank of fractionally integrated time series. It does so by solving a generalized eigenvalue problem of the type proposed by Johansen (1988). However, the limiting distribution of the corresponding trace statistic is χ^2 , where the degrees of freedom depend only on the cointegration rank under the null hypothesis. The usefulness of the asymptotic theory for finite samples is established in a Monte Carlo experiment.

* The first author gratefully acknowledge financial support from the *Sonderforschungsbereich 373* of the DFG. We thank Luis Gil-Alana for helpful comments.

1 Introduction

With his seminal paper introducing fractional integration and cointegration Granger (1981) opened a productive research avenue. Since then cointegration techniques have become standard in the econometrician's tool kit. Fractional cointegration techniques, however, are still to be developed, see Robinson (1994a) and Baillie (1996) for overviews on fractional integration in econometrics. A vector of time series variables is called fractionally cointegrated if the variables are integrated of order $d > 0.5$ and there exists a linear combination of the variables with a smaller degree of integration $d^* < d$. The properties of fractionally cointegrated systems are analyzed by Cheung and Lai (1993), Jeganathan (1999), Robinson and Marinucci (1998), and Tsay (2000). Empirical applications can be found in, e.g., Cheung and Lai (1993), Booth and Tse (1995) and Masih and Masih (1995, 1998), Baillie and Bollerslev (1994) and Dueker and Startz (1998). Recently, Anderson and Gredenhoff (1999) investigated experimentally the behaviour of the likelihood ratio statistic suggested by Johansen (1988). Given observations integrated of order one and fractional cointegration they observe that the likelihood ratio test has a fairly high power. In contrast, Gonzalo and Lee (1998) show that the true null of no cointegration will be rejected spuriously more often than the nominal level given the observed series are fractionally integrated of order $d \neq 1$. This is one motivation for our present paper.

The starting point of our analysis is a univariate score test for integration against fractional alternatives. We propose a regression variant of the score statistic suggested by Robinson (1991, 1994b), Agiakloglou and Newbold (1994), and Tanaka (1999). This test can be understood and implemented as a slight modification of the Dickey-Fuller test (Dickey and Fuller (1979)), although it has a limiting normal distribution. To test against fractional cointegration we generalize the test in the same manner as Johansen's likelihood ratio test for the cointegration rank can be seen as a generalisation of the Dickey-Fuller test. By solving a generalized eigenvalue problem analogous to that by Johansen (1988) we suggest to determine the cointegration

rank of fractionally integrated time series. The limiting distribution depends neither on the order of integration of the series nor on the order of integration of the deviations from the long-run relationships. Instead, the asymptotic distribution is χ^2 , where the degrees of freedom depend on the cointegration rank under the null hypothesis. This result is also valid in case of classical cointegration where the series are integrated of order $d = 1$ with linear combinations integrated of order $d^* = 0$.

The rest of this paper is organized as follows. Section 2 introduces our variant of the univariate score statistic that can be implemented as a modification of the Dickey-Fuller test. Section 3 extends the analysis to multivariate processes to obtain a test for the fractional cointegration rank. The fourth section presents some Monte Carlo evidence on the performance of our tests in finite samples relative to competing procedures. Concluding remarks are found in the final section, and proofs are relegated to the Appendix.

2 Testing against fractional alternatives

Assume that we want to test the hypothesis that a univariate time series is a Gaussian $I(d)$ process, $(1 - L)^d y_t = \varepsilon_t$, $t = 1, 2, \dots, T$, where ε_t is Gaussian white noise with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t^2) = \sigma^2$, against the alternative that y_t is $I(d-b)$. This test is equivalent to a test of the hypothesis that $x_t \equiv (1 - L)^d y_t$ is white noise against the alternative that x_t is $I(-b)$ with $b \neq 0$.

Robinson (1991, 1994b) and Tanaka (1999) derive score statistics for this test problem. Assuming $y_s = 0$ for $s \leq 0$ and normally distributed errors ε_t , the derivative of the log-likelihood function of x_t evaluated at $b = 0$ is obtained as

$$\begin{aligned} \left. \frac{\partial L(b, \sigma)}{\partial b} \right|_{b=0} &= \frac{1}{\sigma^2} \sum_{t=2}^T x_t \left(\sum_{j=1}^{t-1} j^{-1} x_{t-j} \right) \\ &= \frac{1}{\sigma^2} \sum_{t=2}^T x_t x_{t-1}^* , \end{aligned} \tag{1}$$

where

$$x_{t-1}^* = \sum_{j=1}^{t-1} j^{-1} x_{t-j} . \quad (2)$$

Using $\lim_{T \rightarrow \infty} E(x_T^*)^2 = \sigma^2 \pi^2 / 6$, Tanaka (1999) obtains the test statistic

$$\tau_T = \frac{6T}{\pi^2} \left[\sum_{j=1}^{T-1} j^{-1} \hat{\varrho}_j \right]^2 ,$$

where $\hat{\varrho}_j = \sum_{t=j+1}^T x_t x_{t-j} / \sum_{t=1}^T x_t^2$. The same statistic has been proposed by Robinson (1991) to test for no serial correlation against fractionally integrated noise, while Robinson (1994b) favours a frequency domain approximation.

Following Agiakloglou and Newbold (1994) we consider the corresponding regression statistic that is obtained as the squared t -statistic for $\phi = 0$ in the regression¹

$$x_t = \phi x_{t-1}^* + e_t , \quad (3)$$

that is, the test statistic is obtained by using the outer product of gradients as an estimate of the information matrix:

$$\tau_T^* = \frac{\left(\sum_{t=2}^T x_t x_{t-1}^* \right)^2}{\hat{\sigma}_e^2 \sum_{t=2}^T x_{t-1}^{*2}} , \quad (4)$$

where $\hat{\sigma}_e^2$ is the usual estimate for the variance of e_t .

It is interesting to compare the test statistic τ_T^* with the well known Dickey-Fuller test. If $d = 1$ define $x_t = y_t - y_{t-1}$ and $\tilde{x}_{t-1} = y_{t-1} = x_{t-1} + x_{t-2} + \dots + x_1$, and the Dickey-Fuller regression corresponding to (3) is given by

$$x_t = \phi \tilde{x}_{t-1} + e_t .$$

¹Agiakloglou and Newbold (1994) suggest to use $\tilde{x}_{t-1}^{(m)} = \sum_{j=1}^m j^{-1} x_{t-j}$ instead of x_{t-1}^* , where m is some pre-specified truncation parameter. In our test we choose $m = t - 1$ to be time dependent which gives a test statistic that is asymptotically equivalent to the tests suggested by Robinson (1994b) and Tanaka (1999).

The only difference between our statistic in (4) and the Dickey-Fuller statistic is the introduction of the weights j^{-1} in (2). A similar generalization of the Dickey-Fuller test to a fractional framework has recently been proposed by Dolado, Gonzalo and Mayoral (1999). Let $d^* = d - b^*$ be some prespecified value for the fractional parameter. Then Dolado et al. (1999) define $x_t = (1 - L)^d y_t$ and $\tilde{x}_{t-1} = (1 - L)^{d^*} y_{t-1}$, where the latter is given by expansion of the fractional filter. The Dickey-Fuller test results as a special case with $d = 1$ and $d^* = 0$. For $0 \leq d^* < 0.5$ the test has a nonstandard limiting distribution (cf. Dolado et al. 1999). In contrast, as implied by the following lemma, the score statistic (4) is asymptotically χ^2 distributed with one degree of freedom.

Lemma 1: *Let x_t be white noise with $E(x_t) = 0$ and $E(x_t^2) = \sigma^2$. If x_{t-1}^* is generated as in (2) then*

$$(i) \quad T^{-1/2} \sum_{t=2}^T x_t x_{t-1}^* \xrightarrow{d} N(0, \sigma^4 \pi^2 / 6)$$

$$(ii) \quad T^{-1} \sum_{t=2}^T (x_{t-1}^*)^2 \xrightarrow{p} \sigma^2 \pi^2 / 6$$

as $T \rightarrow \infty$.

A similar result was already obtained in Robinson (1994b) and Tanaka (1999) but it is restated here for later reference.

To allow for short memory autocorrelation of x_t we adopt the approach of the REG test suggested by Agiakloglou and Newbold (1994). Assume that x_t is a stable AR(p) process. Then the test statistic builds on residuals $\hat{\varepsilon}_t$ from a regression of x_t on x_{t-1}, \dots, x_{t-p} instead of the original series. Furthermore, x_{t-1}^* is replaced by $\hat{\varepsilon}_{t-1}^*$ computed like in (2) and by lagged values of x_t , that is, x_{t-1}^* is replaced by $\tilde{\varepsilon}_{t-1}^* = [\hat{\varepsilon}_{t-1}^*, x_{t-1}, \dots, x_{t-p}]'$. The use of $\hat{\varepsilon}_t$ and $\tilde{\varepsilon}_{t-1}^*$ in (4) results again in a limiting χ^2 distribution with one degree of freedom.

Furthermore, the series x_t can be demeaned or detrended prior to computing τ_T^* . Following Robinson (1994b) this does not affect the limiting distribution.

As Johansen's (1988) test for the cointegration rank can be seen as the multivariate extension of the Dickey-Fuller test, a multivariate version of the score test given in (4) can be constructed to test for the cointegration rank of a vector of time series.

3 Testing hypotheses on the cointegration rank

First, we neglect the possibility of a cointegration relationship. Assume that y_t is an $n \times 1$ vector of $I(d)$ components and under the alternative hypothesis it is assumed that all components of y_t are $I(d - b)$. The assumption that d and b are the same for all components of y_t is made to keep the notation reasonably simple and will be relaxed below (see Remark C).

As in the univariate case the score is obtained as

$$\begin{aligned} \left. \frac{\partial L(b, \Sigma)}{\partial b} \right|_{b=0} &= \sum_{t=2}^T x_t' \Sigma^{-1} \left(\sum_{j=1}^{t-1} j^{-1} x_{t-j} \right) \\ &= \sum_{t=2}^T x_t' \Sigma^{-1} x_{t-1}^* , \end{aligned}$$

or

$$\left. \frac{\partial L(b, \Sigma)}{\partial b} \right|_{b=0} = \text{tr}[\Sigma^{-1} S_{10}] , \quad (5)$$

where $x_t = (1 - L)^d y_t$, $\Sigma = E(x_t x_t')$, and

$$S_{10} = \sum_{t=2}^T x_{t-1}^* x_t' \quad \text{with} \quad x_{t-1}^* = \sum_{j=1}^{t-1} j^{-1} x_{t-j} . \quad (6)$$

If Σ is unknown, it is replaced by a consistent estimate denoted as $\hat{\Sigma}$. This gives rise to the multivariate score test statistic

$$\Lambda_0(d) = \text{tr}[\hat{\Sigma}^{-1} S_{10}' S_{11}^{-1} S_{10}] , \quad (7)$$

where

$$\hat{\Sigma} = T^{-1} \sum_{t=1}^T x_t x_t' , \quad S_{11} = \sum_{t=2}^T x_{t-1}^* x_{t-1}^{*'} .$$

Since the trace is the sum of eigenvalues it is useful to consider the eigenvalues of the problem

$$|\lambda \widehat{\Sigma} - S'_{10} S_{11}^{-1} S_{10}| = 0$$

that are given by

$$\lambda_j = \frac{v'_j S'_{10} S_{11}^{-1} S_{10} v_j}{v'_j \widehat{\Sigma} v_j} \quad j = 1, \dots, n,$$

where v_j is the corresponding eigenvector. It turns out that λ_j is the test statistic for $\phi = 0$ in the auxiliary regression

$$(v'_j x_t) = \phi' x_{t-1}^* + e_t.$$

Using a multivariate version of Lemma 1, it is straightforward to show that λ_j is asymptotically χ^2 distributed with n degrees of freedom. Furthermore, since the eigenvalues are asymptotically uncorrelated, we obtain

$$\Lambda_0(d) \xrightarrow{d} \chi^2(n^2).$$

So far we have considered the hypothesis that all possible linear combinations of y_t have the same fractional parameter d . In what follows we consider the hypothesis that y_t is *fractionally cointegrated* as specified in the following definition, which is adapted from Engle and Granger (1987).

Definition 1: *The $n \times 1$ vector y_t is cointegrated of order $CI(d, b)$ and rank $0 < r < n$ if there exist a $n \times r$ cointegration matrix β of full rank and a linearly independent $n \times (n - r)$ matrix γ such that*

$$\begin{aligned} \gamma' y_t &\sim I(d) \\ \beta' y_t &\sim I(d - b) \end{aligned}$$

where $d > 0.5$ and $b > 0$.

This definition is similar to the concept used by Cheung and Lai (1993) and Dueker and Startz (1998), for example. In what follows we assume that the fractional parameter d is given, while b is assumed to be unknown.

For the ease of exposition we first make the following assumption.

Assumption 1: Let $y_t = (1 - L)^{-d}x_t$ be $CI(d, b)$ and

$$u_t = \begin{bmatrix} \gamma'x_t \\ (1 - L)^{-b}\beta'x_t \end{bmatrix}$$

is white noise with $E(u_t|u_{t-1}, u_{t-2}, \dots) = 0$ and a positive definite covariance matrix $\Omega = E(u_t u_t')$.

The assumption that u_t is white noise is made to facilitate the proof of the main result, and it will be relaxed below (see Remark B). For $d = 1$ and $b = 1$ this assumption is equivalent to assuming that y_t can be represented by a VAR(1) system.

Following Johansen (1995) we test hypotheses on the cointegration rank based on the sum of the $n - r$ smallest eigenvalues of the problem:

$$|\lambda S_{11} - S_{10} \widehat{\Sigma}^{-1} S_{10}'| = 0 ,$$

or, equivalently,

$$|\lambda \widehat{\Sigma} - S_{10}' S_{11}^{-1} S_{10}| = 0 , \quad (8)$$

where $\widehat{\Sigma}$, S_{10} and S_{11} are defined as in (7). As in the univariate case, we replace the partial sum $y_{t-1} = x_1 + \dots + x_{t-1}$ used in Johansen (1988) by the weighted sum $x_{t-1}^* = \sum_{j=1}^{t-1} j^{-1} x_{t-j}$. Furthermore, for the special case $r = 0$ the corresponding trace statistic turns out to be identical to the multivariate score statistic $\Lambda_0(d)$ defined in (7).

In the following theorem, the null distribution of the test statistic for the cointegration rank is given.

Theorem 1: Let y_t be $CI(d, b)$ with $d > 0.5$ and $b > 0$. Under Assumption 1 and the hypothesis $r = r_0$ the test statistic

$$\Lambda_{r_0}(d) = \sum_{j=1}^{n-r_0} \lambda_j ,$$

where $\lambda_1 \leq \dots \leq \lambda_n$ are the ordered eigenvalues of problem (8), has an asymptotic χ^2 distribution with $(n - r_0)^2$ degrees of freedom. Under the alternative $H_1: r_0 < r$ the test statistic diverges to infinity at rate T .

REMARK A: To allow for a non-zero mean we assume $E(x_t) = \theta' d_t$, where d_t is a $k \times 1$ vector of deterministic functions like a constant, a time trend or dummy variables, and θ is a $k \times n$ matrix of parameters. Since $x_t = (1 - L)^d y_t$ is assumed to be stationary, the least-squares regression of x_t on d_t yields a \sqrt{T} -consistent estimate of θ . The test statistic can be constructed by using the adjusted series $\tilde{x}_t = x_t - \hat{\theta}' d_t$ instead of the original observations x_t . Following Robinson (1994b) it can be shown that the limiting distribution is not affected when the time series is adjusted for deterministic terms like a constant or a linear time trend.

REMARK B: To allow for possible short run dynamics, the approach of Agiakloglou and Newbold (1994) can be adapted. Assume that $\gamma' x_t$ has a VAR(p) representation.² In this case the test statistic is constructed by using the prewhitened series, i.e., the residuals $\hat{\varepsilon}_t$ from a regression of x_t on x_{t-1}, \dots, x_{t-p} instead of the original series. Furthermore, the vector x_{t-1}^* is replaced by $\hat{\varepsilon}_{t-1}^*$ and lagged values of x_t , where $\hat{\varepsilon}_{t-1}^*$ computed like in (2). More precisely, x_{t-1}^* is replaced by $\tilde{\varepsilon}_{t-1}^* = [\hat{\varepsilon}_{t-1}^{*'}, x'_{t-1}, \dots, x'_{t-p}]'$. Using $\hat{\varepsilon}_t$ and $\tilde{\varepsilon}_{t-1}^*$ when computing the matrices in (8), the resulting test statistic given in Theorem 1 has the same asymptotic χ^2 distribution with $(n - r)^2$ degrees of freedom.

REMARK C: The test procedure can be straightforwardly extended to allow for different values of d_1, \dots, d_n and b_1, \dots, b_r . In this case we define cointegration among the vector y_t if there exist $r > 0$ linear combinations that are integrated of order $[\min(d_1, \dots, d_n) - b_k]$, where $b_k > 0$ and $k = 1, \dots, r$. This definition may appear inappropriate as one usually assumes that the

²In the proof of Theorem 1 it is shown that asymptotically the test statistic does not depend on the component $\beta' x_t$. Therefore, we do not need to specify the short run dynamics for $\beta' x_t$.

variables in a cointegration relationship have the same order of integration. However, from the literature on $I(2)$ cointegrated variables it is known that series with a different order of integration may constitute a cointegration relationship whenever the higher order integrated components cancel out. For example assume that the first variable y_{1t} is $I(0.8)$, the second variable y_{2t} is $I(0.6)$ and the third variable is given by $y_{3t} = y_{1t} + y_{2t} + u_t$, where u_t is $I(0.2)$. Then y_{3t} is $I(0.8)$ and the vector $\beta = [-1, -1, 1]'$ yields a cointegration relationship with $b_1 = 0.4$.

To apply the test procedure to this case, we construct the components of the vector x_t as $x_{kt} = (1 - L)^{d_k} y_{kt}$ and the vector x_{t-1}^* is computed as in (6). Since the null distribution of the test statistic does not depend on b_1, \dots, b_r , the results of Theorem 1 are also valid for the case of different values of b_1, \dots, b_r .

4 Finite sample properties

Since our multivariate test is based on a generalization of the regression variant of the score statistic, it is interesting to compare the performance of our variant to other univariate tests against fractional alternatives. We consider the two-sided test problem with the null hypothesis $d = 1$ against the alternative $d_1 = d - b \neq 1$.

For our Monte Carlo experiments we simulated fractionally integrated noise according to Hosking (1984),

$$(1 - L)^{1-d_1} x_t = \varepsilon_t, \quad \varepsilon_t \sim iid\mathcal{N}(0, 1), \quad \text{and } 0.5 < d_1 < 1.5. \quad (9)$$

The final series are obtained by computing the partial sum $y_t = \sum_{j=1}^t x_j$.

The Monte Carlo comparison includes Robinson's (1994b) frequency domain score statistic R_T , Tanaka's (1999) time domain statistic τ_T , the REG(m) statistic by Agiakloglou and Newbold (1994) relying on $\tilde{x}_{t-1}^{(m)} = \sum_{j=1}^m j^{-1} x_{t-j}$ instead of x_{t-1}^* , and our variant τ_T^* defined in (4). With the REG test we tried $m = 10$ and $m = 20$ and observed only marginal differences. Here, only the slightly more powerful case $m = 10$ is reported. Furthermore, the fractional

Dickey-Fuller (FDF) test suggested by Dolado, Gonzalo and Mayoral (1999) is applied using $d^* = d_1$ under the alternative, and $d^* = 0.9$ under the null hypothesis. All tests are computed by using demeaned series x_t .

For the two-sided test problem all statistics are asymptotically χ^2 distributed with one degree of freedom. In Table 1 we report the rejection frequencies using a nominal significance level of 0.05. All results rely on 5000 Monte Carlo replications.

The results presented in Table 1 can be summarized as follows. The actual sizes of R_T and τ_T are slightly below the nominal size, while the other tests are only slightly above. For $T = 250$, REG(10) and τ_T^* are closest to the nominal level. Our findings with respect to the power are as follows. For $T = 100$ and $d_1 < 1$ our Dickey-Fuller type score test dominates all other tests. The REG(10) test is only slightly less powerful. For $T = 100$ and $d_1 > 1$, Tanaka's (1999) test performs best, the other score tests behave more or less the same, and the FDF test of Dolado et al. (1999) has slightly less power (except for $d_1 = 1.1$). For $T = 250$ and $d_1 < 1$, τ_T^* outperforms again all other tests, for $T = 250$ and $d_1 > 1$, the performance of all score type tests is very similar. The fact that τ_T^* from (4) is most powerful against $d_1 < 1$ makes our test a promising candidate for testing against cointegration.

Next, in a bivariate framework the performance of the test against fractional alternatives is compared to the LR test statistic suggested by Johansen (1988). The data is generated according to the model

$$\begin{aligned} y_{1t} &= \alpha y_{1,t-1} + u_t, & u_t &= \rho u_{t-1} + \varepsilon_{1t}, & |\rho| &< 1 \\ y_{2t} &= y_{1t} + v_t, & (1-L)^{1-b} v_t &= \varepsilon_{2t}, \end{aligned} \quad (10)$$

where ε_{1t} and ε_{2t} are uncorrelated white noise processes distributed as $N(0, 1)$. If $\alpha = 1$ and $b > 0$, then the process has a cointegration relationship with cointegration vector $[1, -1]$. If $\alpha = 0$, then y_{1t} is stationary and if in addition $b = 0$, then the cointegration rank is one with the ‘‘cointegration vector’’ $[1, 0]$, whereas for $b > 0$ the cointegration rank $r = 2$ results. Furthermore, for $\rho = 0$ the process has a VAR(1) representation, whereas for $\rho \neq 0$ a VAR(2) representation is required. In the latter case Johansen's LR test includes a lagged difference and the statistic $\Lambda_r(1)$ from Theorem 1 against

fractional alternatives (FRAC) is based on the prewhitened series resulting from the residuals of a VAR(1) model for the first differences (see Remark B). The sample size is $T = 100$ and 5000 Monte Carlo replications are used to compute the rejection frequencies. A nominal significance level of 0.05 is used for all experiments.

Table 2 presents the results of the Monte Carlo simulations. It turns out that for all tests the actual sizes are close to the nominal ones. Under the alternative ($0 < b \leq 0.5$) the FRAC test outperforms Johansen's LR test against $b = 1$. However the gain in power seems to depend on the short-run dynamics. If $r = 1$ and $\rho = 0.5$, then the power of the FRAC test is roughly similar to Johansen's test, whereas for $\rho = 0$ and $\rho = -0.5$ the FRAC test is clearly superior.

5 Concluding remarks

We modified the efficient score tests by Robinson (1991, 1994b), Agiakloglou and Newbold (1994) and Tanaka (1999) for the fractional order of integration of univariate time series in such a way that the statistic can be computed from a Dickey-Fuller type regression. The test relies on a limiting χ^2 distribution with one degree of freedom.

In this paper we suggest a straightforward extension of the univariate test to a multivariate setup. It allows to determine the cointegration rank of possibly fractionally integrated series, where the cointegrated linear combinations may be fractionally integrated as well. Just like Johansen's (1988) LR test statistic can be seen as a generalization of the univariate Dickey-Fuller test, our test is a multivariate version of the regression based score statistic.

In the multivariate framework the statistic builds on eigenvalues from a generalized eigenvalue problem, however, under the null hypothesis of r linearly independent cointegrating vectors the limiting distribution is χ^2 . The degrees of freedom depend only on the cointegration rank under the null hypothesis and are affected neither by the potentially fractional order of integration of the series nor by the possibly fractional order of integration

for the deviations of the long-run relationship.

Although the new procedure is quite general in that it allows for fractional integration and cointegration, our test is based on a severe restriction. Our test requires that all series are integrated of some known order d . One possibility to relax this assumption is to compute the test statistic with an estimated order d and use bootstrap critical values (see Davidson 2000).

Appendix

Proof of Lemma 1:

Using $\sum_{j=1}^{\infty} j^{-2} = \pi^2/6$ we have

$$\lim_{t \rightarrow \infty} E(x_t^*)^2 = \sigma^2 \pi^2/6 .$$

Furthermore, $v_t = x_t x_{t-1}^*$ is a martingale difference sequence with $E(v_t | x_{t-1}, x_{t-2}, \dots) = 0$ and $E(v_t^2) \leq \sigma^4 \pi^2/6$. Using the central limit theorem for martingale difference sequences (e.g. White, 1984), it follows that $T^{-1/2} \sum v_t$ converges to a normal distributed random variable with expectation zero and variance $\sigma^4 \pi^2/6$. Under similar conditions, the law of large number implies that $T^{-1} \sum_{t=2}^T (x_{t-1}^*)^2$ converge in probability to $\sigma^2 \pi^2/6$. ■

Proof of Theorem 1:

Let γ be a $n \times (n - r)$ matrix such that $\gamma' x_t$ is white noise. Furthermore we multiply $\gamma' x_t$ by the Choleski factor of the inverse of $\gamma' \Sigma \gamma$ such that $z_{1t} = (\gamma' \Sigma \gamma)^{-1/2} \gamma' x_t$ has a unit covariance matrix. Under the null hypothesis, there exists a matrix β such that $\beta' x_t$ is $I(-b)$. We define

$$\tilde{z}_{2t} = \beta' x_t - \beta' \Sigma \gamma (\gamma' \Sigma \gamma)^{-1} \gamma' x_t$$

so that \tilde{z}_{2t} is uncorrelated with z_{1t} . Furthermore, we multiply \tilde{z}_{2t} by the Choleski factor of the inverse of $\tilde{\Sigma}_2 = E(\tilde{z}_{2t} \tilde{z}_{2t}')$ so that $z_{2t} = \tilde{\Sigma}_2^{-1/2} \tilde{z}_{2t}$ has a unit covariance matrix.

It is convenient to define the matrices $Z_1 = [z_{12}, \dots, z_{1T}]'$, $Z_2 = [z_{22}, \dots, z_{2T}]'$, $Z_1^* = [z_{11}^*, \dots, z_{1,T-1}^*]'$, $Z_2^* = [z_{21}^*, \dots, z_{2,T-1}^*]'$, and $Z = [Z_1, Z_2]$, where

$$z_{kt}^* = \sum_{j=1}^{t-1} j^{-1} z_{k,t-j} \quad \text{for } k = 1, 2.$$

The following Lemma presents some asymptotic properties of the moments that will be used in what follows.

Lemma A.1: Let Z_{ki} and Z_{ki}^* denote the i 'th columns of Z_k and Z_k^* ($k = 1, 2$). As $T \rightarrow \infty$ we have

$$\begin{aligned}
(i) \quad & T^{-1/2} Z'_{1i} Z^*_{1j} \Rightarrow N(0, \pi^2/6) & (ii) \quad & T^{-1/2} Z'_{1i} Z_{2j} \Rightarrow N(0, 1) \\
(iii) \quad & T^{-1/2} Z'_{1i} Z^*_{2j} \Rightarrow N(0, v_{ij}) & (iv) \quad & T^{-1} Z'_{2i} Z^*_{1j} \xrightarrow{p} \tau_{ij} \\
(v) \quad & T^{-1} Z'^*_{1i} Z^*_{1i} \xrightarrow{p} \pi^2/6 & (vi) \quad & T^{-1} Z'_{2i} Z^*_{2j} \xrightarrow{p} \rho_{ij} \\
(vii) \quad & T^{-1} Z'^*_{2i} Z^*_{2i} \xrightarrow{p} v_{ii} & (viii) \quad & T^{-1} Z'^*_{1i} Z^*_{2i} \xrightarrow{p} c_{ij}
\end{aligned}$$

where τ_{ij} , ρ_{ij} , v_{ij} and c_{ij} are the (i, j) elements of the matrices $E(z_{2t} z_{1t}^{*'})$, $E(z_{2t} z_{2t}^{*'})$, $E(z_{2t}^* z_{2t}^{*'})$ and $E(z_{1t}^* z_{2t}^{*'})$, respectively.

PROOF: (i) follows directly from Lemma 1. (ii) follows from the fact that z_{1t} is white noise and uncorrelated with z_{2t} . (iii) and (iv) can be shown by using an appropriate version of the central limit theorem. Since the asymptotic normality of these terms is not important for our results we do not give the results for the asymptotic variances. (v) follows from Lemma 1 and (vi) – (viii) can be shown by using the law of large numbers.

Now consider the eigenvalue problem (8) which is equivalent to

$$|\lambda T^{-1} Z' Z - Z' Z^* (Z^{*'} Z^*)^{-1} Z^{*'} Z| = 0$$

Using the results of Lemma A.1 it is not difficult to verify that

$$Z' Z^* (Z^{*'} Z^*)^{-1} Z^{*'} Z = \begin{bmatrix} O_p(1) & O_p(T^{1/2}) \\ O_p(T^{1/2}) & O_p(T) \end{bmatrix}.$$

It follows from rule (6) of Sec. 5.3.1 in Lütkepohl (1996) that $n - r$ eigenvalues are $O_p(1)$ and the other r eigenvalues are $O_p(T)$. Let $v^* = [v_1^*, \dots, v_r^*]' = [I_r, \Phi']'$ denote the re-normalized eigenvectors resulting from $v^* = v(v_{(1)})^{-1}$, where $v_{(1)}$ is the upper $r \times r$ block of the original matrix of eigenvectors v . As the original eigenvectors, the re-normalized eigenvectors fulfill the equation

$$[\lambda T^{-1} Z' Z - Z' Z^* (Z^{*'} Z^*)^{-1} Z^{*'} Z] v_j^* = 0$$

where $v_j^* = [v_j', \phi_j']'$ and v_j is the j 'th column of I_r and ϕ_j is the j 'th column of Φ . From the lower $n - r$ equations we obtain

$$\lambda T^{-1} Z'_2 (Z_{(j)} - Z_2 \phi_j) - Z'_2 Z^* (Z^{*'} Z^*)^{-1} Z^{*'} (Z_{(j)} - Z_2 \phi_j) = 0, \quad (11)$$

where $Z_{(j)}$ denotes the j 'th column of Z . Since for $j = 1, \dots, n - r$ the first term in (11) is $O_p(1)$ and the second term is $O_p(T)$, it follows that

$$Z_2' Z^* (Z^{*'} Z^*)^{-1} Z^{*'} (Z_{(j)} - Z_2 \phi_j) = O_p(1) .$$

Hence,

$$\phi_j = [Z_2' Z^* (Z^{*'} Z^*)^{-1} Z^{*'} Z_2]^{-1} Z_2' Z^* (Z^{*'} Z^*)^{-1} Z^{*'} Z_{(j)} + O_p(T^{-1}) .$$

It follows that ϕ_j is asymptotically equivalent to a 2SLS estimator of ϕ_j^* in the equation $Z_{(j)} = Z_2 \phi_j^* + e_j$ using the instrumental matrix Z^* . Define

$$\hat{e}_j = Z_{(j)} - Z_2 \phi_j^* ,$$

then the respective eigenvalue results as

$$\lambda_j = \frac{\hat{e}_j' Z^* (Z^{*'} Z^*)^{-1} Z^{*'} \hat{e}_j}{T^{-1} \hat{e}_j' \hat{e}_j} + o_p(1) .$$

This shows that the eigenvalue is asymptotically equivalent to Sargan's (1958) test of over-identifying restrictions. It follows that the denominator can be re-written as

$$Z_{(j)}' M_Z^* Z_{(j)}$$

where

$$M_Z^* = \{ Z^* (Z^{*'} Z^*)^{-1} Z^{*'} - Z^* (Z^{*'} Z^*)^{-1} Z^{*'} Z_2 [Z_2' Z^* (Z^{*'} Z^*)^{-1} Z^{*'} Z_2]^{-1} Z_2' Z^* (Z^{*'} Z^*)^{-1} Z^{*'} \} .$$

Since the elements of $Z_{(j)}$ are assumed to be uncorrelated with unit variance and M_Z^* is an idempotent matrix with rank $n - r$, it follows that λ_j is asymptotically χ^2 distributed with $n - r$ degrees of freedom (see also Hansen, 1982).

Furthermore, since \hat{e}_j is asymptotically independent of \hat{e}_k , the eigenvalues are independently distributed as $T \rightarrow \infty$ and, thus, $\sum_{j=1}^{n-r} \lambda_j$ is asymptotically χ^2 distributed with $(n - r)^2$ degrees of freedom.

Under the alternative \hat{e}_j and Z^* are correlated so that

$$\lim_{T \rightarrow \infty} E(T^{-1} \hat{e}_j' Z^*) = \mu_{eZ}^* \neq 0 .$$

Thus, the test statistic is $O_p(T)$ under the alternative. ■

References

- Agiakloglou, Ch., and Newbold, P.** (1994), Lagrange Multiplier Tests for Fractional Difference, *Journal of Time Series Analysis* 15, 253-262.
- Andersson, M.K., and Gredenhoff, M.P.** (1999), On the Maximum Likelihood Cointegration Procedure under a Fractional Equilibrium Error, *Economics Letters* 65, 143-147.
- Baillie, R.T.** (1996), Long Memory Processes and Fractional Integration in Econometrics, *Journal of Econometrics* 73, 5-59.
- Baillie, R.T., and Bollerslev, T.** (1994), Cointegration, Fractional Cointegration, and Exchange Rate Dynamics, *The Journal of Finance* 49, 737-745.
- Booth, G.G., and Tse, Y.** (1995), Long Memory in Interest Rate Future Markets: A Fractional Cointegration Analysis, *The Journal of Future Markets* 15, 563-54.
- Cheung, Y.-E., and Lai, K.S.** (1993). A Fractional Cointegration Analysis of Purchasing Power Parity, *Journal of Business and Economic Statistics* 11, 103-122.
- Davidson, J.** (2000), A Model of Fractional Cointegration, and Tests for Cointegration Using the Bootstrap, Cardiff University, mimeo.
- Dickey, D.A., and Fuller, W.A.** (1979), Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association* 74, 427-431.
- Dolado, J.J., Gonzalo, J., and Mayoral, L.** (1999), A Fractional Dickey-Fuller Test, mimeo Universidad Carlos III de Madrid.
- Dueker, M., and Startz, R.** (1998), Maximum-Likelihood Estimation of Fractional Cointegration with an Application to U.S. and Canadian Bond Rates, *The Review of Economics and Statistics* 80, 420-426.

- Engle, R.F. and C.W.J. Granger, C.W.J.** (1987), Co-integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, 55, 251–276.
- Gonzalo, J., and Lee, H.-T.** (1998), Pitfalls in Testing for Long Run Relationships, *Journal of Econometrics* 86, 129-154.
- Granger, C.W.J.** (1981), Some Properties of Time Series Data and their Use in Econometric Model Specification, *Journal of Econometrics* 16, 121-130.
- Hansen, L.P.** (1982), Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica* 50, 1029-1054.
- Hosking, J.R.M.** (1984), Modeling Persistence in Hydrological Time Series using Fractional Differencing, *Water Resources Research* 20, 1898-1908.
- Jeganathan, P.** (1999), On Asymptotic Inference in Cointegrated Time Series with Fractionally Integrated Errors, *Econometric Theory* 15, 583-621.
- Johansen, S.** (1988), Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control* 12, 231-254.
- Johansen, S.** (1995), Likelihood-Based Inference in Cointegrated Vector Autoregressive Models, Oxford University Press.
- Lütkepohl** (1996), *Handbook of Matrices*, John Wiley.
- Masih, R. and Masih, A.M.M.** (1995), A Fractional Cointegration Approach to Empirical Tests of PPP: New Evidence and Methodological Implications from an Application to the Taiwan/US Dollar Relationship, *Review of World Economics* 131, 673-694.
- Masih, A.M.M. and Masih, R.** (1998), A Fractional Cointegration Approach to Testing Mean Reversion Between Spot and Forward Ex-

change Rates: A Case of High Frequency Data with Low Frequency Dynamics, *Journal of Business Finance and Accounting* 25, 987-1003.

Robinson, P.M. (1991), Testing for Strong Serial Correlation and Dynamic Conditional Heteroskedasticity in Multiple Regressions, *Journal of Econometrics* 47, 67-84.

Robinson, P.M. (1994a), Time Series with Strong Dependence, in C.A. Sims (ed.), *Advances in Econometrics. Sixth World Congress, Vol. I*, Cambridge University Press, Cambridge, 47-95.

Robinson, P.M. (1994b), Efficient Tests of Nonstationary Hypotheses, *Journal of the American Statistical Association* 89, 1420-1437.

Robinson, P.M., and Marinucci, D. (1998), Semiparametric Frequency Domain Analysis of Fractional Cointegration, STICERD working paper, EM/98/350, LSE.

Sargan, J.D. (1958), The Estimation of Economic Relationships Using Instrumental Variables, *Econometrica*, 26, 393-415.

Tanaka, K. (1999), The Nonstationary Fractional Unit Root, *Econometric Theory* 15, 549-582.

Tsay, W.-J. (2000), Estimating Trending Variables in the Presence of Fractionally Integrated Errors, *Econometric Theory* 16, 324-346.

White, H. (1984), *Asymptotic Theory for Econometricians*, Academic Press.

Table 1: Size and power of the univariate tests

$T = 100$					
d_1	R_T	τ_T	Reg(10)	τ_T^*	FDF
0.7	0.536	0.771	0.886	0.937	0.879
0.8	0.304	0.417	0.603	0.695	0.532
0.9	0.098	0.109	0.221	0.275	0.169
1.0	0.033	0.024	0.057	0.066	0.068
1.1	0.133	0.141	0.132	0.116	0.144
1.2	0.485	0.515	0.475	0.456	0.429
1.3	0.810	0.835	0.801	0.797	0.711
$T = 250$					
d_1	R_T	τ_T	REG(10)	τ_T^*	FDF
0.7	0.998	0.999	0.999	1.000	0.999
0.8	0.903	0.932	0.955	0.974	0.910
0.9	0.345	0.373	0.470	0.533	0.351
1.0	0.042	0.038	0.052	0.053	0.065
1.1	0.401	0.407	0.400	0.353	0.317
1.2	0.925	0.929	0.922	0.914	0.815
1.3	0.999	0.999	0.999	0.999	0.982

Note: Two-sided tests for the null hypothesis $d = 1$. The rejection frequencies are based on 5000 replications of the model (9). R_T is Robinson's (1994b) frequency domain statistic, τ_T denotes Tanaka's score statistic, REG(10) denotes the regression test of Agiakloglou and Newbold (1994) with truncation lag $m = 10$, τ_T^* is the score statistic given in (4), and FDF denotes the statistic suggested by Dolado et al. (1999). The nominal size for all tests is 0.05.

Table 2: Size and power against fractional alternatives

$H_0 : r = 0 \quad (\alpha = 1)$						
	$\rho = 0$		$\rho = 0.5$		$\rho = -0.5$	
b	FRAC	LR	FRAC	LR	FRAC	LR
0.0	0.045	0.041	0.065	0.050	0.056	0.045
0.1	0.108	0.065	0.095	0.063	0.085	0.056
0.2	0.349	0.159	0.176	0.109	0.169	0.099
0.3	0.695	0.400	0.322	0.318	0.323	0.206
0.5	0.985	0.909	0.731	0.657	0.999	0.983
$H_0 : r = 1 \quad (\alpha = 0)$						
	$\rho = 0$		$\rho = 0.5$		$\rho = -0.5$	
b	FRAC	LR	FRAC	LR	FRAC	LR
0.0	0.044	0.054	0.044	0.061	0.050	0.060
0.1	0.191	0.150	0.109	0.108	0.124	0.109
0.2	0.594	0.347	0.280	0.245	0.307	0.245
0.3	0.909	0.622	0.525	0.434	0.577	0.419
0.5	0.999	0.982	0.886	0.873	0.935	0.879

Note: Entries of the table report the rejection frequencies of the tests for the null hypothesis $r = r_0$ ($b = 0$). FRAC is the test suggested in Theorem 1 and LR denotes Johansen's trace statistic. The results are based on 5000 replications of model (10) with $T = 100$. The nominal size is 0.05. For $\rho = 0$ the test are based on a VAR(1) model and for $\rho \neq 0$ on a VAR(2) representation.